Holiday or vacation? The processing of variation in vocabulary across dialects

Martin, Clara D.; Garcia, Xavier; Potter, Douglas; Melinger, Alissa; Costa, Albert

Published in:
Language, Cognition and Neuroscience

DOI:
10.1080/23273798.2015.1100750

Publication date:
2016

Holiday or *vacation*? The processing of variation in vocabulary across dialects.

Running head: Dialectal variation in speech perception

Clara D. Martin [a,b], Xavier Garcia [c], Douglas Potter [d], Alissa Melinger [d] & Albert Costa [c,e]

a. BCBL. Basque Center on Cognition, Brain and Language, Paseo Mikeletegi, 69, 20009 Donostia-San Sebastian, Spain

b. IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Spain

c. University Pompeu Fabra, Departament de Tecnologies de la Informaciò i les Comunicacions, carrer Roc Boronat, 138, 08018 Barcelona, Spain

d. University of Dundee, School of Psychology, Dundee, Angus DD1 4HN, United Kingdom

e. ICREA, Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

Corresponding Author

Dr. Clara D. Martin
Basque Center on Cognition, Brain and Language
Parque Tecnológico de Miramón
Paseo Mikeletegi 69, 2nd
20009 San Sebastian
Spain
Work Phone: 00 34 943 309 300 (ext: 204)
Fax: 00 34 943 309 052
E-mail: c.martin@bcbl.eu

Abstract

Native speakers with different linguistic backgrounds differ in their usage of language, and particularly in their vocabulary. For instance, British natives would use the word '*holiday*' when American natives would prefer the word '*vacation*'. This study investigates how cross-dialectal lexical variation impacts lexical processing, by using event-related potentials. Electrophysiological responses were recorded while British natives listened to British or American speech in which lexical frequency dominance across dialects was manipulated (British *versus* American vocabulary). Words inconsistent with the dialect of the speaker (British words uttered by American speakers and vice versa) elicited larger negative electrophysiological deflections than consistent words, around 700 ms after stimulus onset. Thus, processing of British words was easier when listening to British speakers and processing of American words was easier when listening to American speakers. These results show that listeners integrate their knowledge about cross-dialectal lexical variations in vocabulary as speech unfolds, as it was previously shown for social lexical variations. Moreover, it shows that speech perception is influenced by speaker's attributes not only at the conceptual but also at the lexical level. These results have important methodological implications, revealing that context-dependent information (such as speaker's accent) better explains lexical integration than context-independent lexical frequency.

Keywords

Language perception; Sociolinguistics; Event-Related Potentials; Dialect comprehension; lexical integration

Introduction

Speakers with different social and linguistic backgrounds differ in their usage of language, a phenomenon named sociolinguistic variation (Labov, 1972). For example, adults use the word *wine* more frequently than children, older speakers use the word *frightened* more than younger speakers, and females use the word *husband* more frequently than males (see van Berkum et al., 2008; Walker & Hay, 2011). Other examples of this variability are the differences in vocabulary between speakers of different regional dialects, such as British and American English (e.g.; *holiday* is more frequent than *vacation* in British than in American English, and vice versa). Thus, British natives may watch an advertisement for a perfect *holiday* plan on a British channel, whereas on American satellite channels the same advertisement would be announced as a perfect *vacation* plan. Listeners have to deal with these variable linguistic forms to achieve comprehension, even if the message is conveyed with infrequent forms that are common in other groups of people. Parsers are not blind to these differences and acquire knowledge about sociolinguistic variation through their experience with speakers from different communities (Guy, 2011; Barsalou, 2012). The current experiment addresses the question of whether listeners adjust their vocabulary preferences according to what they know about the dialect of the speaker. Particularly, we investigated how lexical processing is modulated by information about the speaker’s dialect, inferable from his accent. In other words, we explored whether British natives listening to American speakers find it easier to integrate the British word "*holiday*", because it is the most frequent in their vocabulary, or the American word "*vacation*", because it is the most frequent in the speaker’s dialect.

Sociolinguistic variation is a broad term and can refer to several different types of alternations. First, we will use the term cross-dialectal lexical variations to refer to two synonymous lexical entries, one being used more frequently in one dialect than another, and vice versa. For instance, British speakers are more likely to use the word *barrister* while American speakers are more likely to use the synonymous alternative *attorney* (Walker & Hay, 2011). This type of sociolinguistic variation will be the focus of the present study. Before focusing on cross-dialectal lexical variations, we present previous literature on other types of sociolinguistic variations: By sociological lexical variations, we refer to different lexical entries, one being more frequently used by one sociological group than another, and vice versa. For instance, the word *husband* is used more frequently by women and the word *wife* by men (van Berkum et al., 2008; Walker & Hay, 2011). By sociological phonetic variations, we refer to two phonetic/phonological forms of a given word, one being more frequently encountered in one sociological group than another, and vice versa. For instance, vowel fundamental and formant frequencies are in general lower for male than female talkers (see Strand, 1999). Finally, we will use the term cross-dialectal phonetic variations to refer to different phonetic/phonological forms of a given word, one being more frequently encountered in one dialect than another, and vice versa. For instance, '-er' final words in English are pronounced with a final rhotic /r/ in Standard American English, but not in Standard British English (Walker & Hay, 2011; see also Sumner & Samuel, 2009 for variations across American dialects).

Previous research on sociolinguistic variation already showed extensively that sociological and cross-dialectal phonetic variations influence speech processing (see Drager, 2010 for a review). The main contribution of this important research topic has been to show that listeners retune their language perception according to their knowledge and experience with the speaker's normal usage of language (Niedzielski, 1999; Evans & Iverson, 2004; Johnson, 2006; Foulkes & Docherty, 2006; Sumner & Samuel, 2009; Campbell-Kibler, 2010; Creel & Bregman, 2011; Goslin et al., 2012). Regarding sociophonetic variations, several studies have shown that listeners perceive speech sounds differently depending on the apparent gender / social characteristics of the speaker (Remez et al., 1987; Strand, 1999; Johnson et al., 1999; Johnson, 2005; Hay et al., 2006; see Drager, 2010; Thomas, 2002 for reviews). In eye-tracking studies, it was shown that knowledge about the speaker (voice characteristics) can prompt expectations for visual targets before those targets are actually mentioned in speech (Creel, 2012; Horton & Slaten, 2012). Cross-dialectal phonetic variations were also extensively studied, showing that phonetic variants consistent with the accent of the speaker are easier to process than phonetic variants that are inconsistent with the speaker's accent. This has been shown, for instance, by Loudermilk and colleagues (2013a,b), investigating the realization of the English morpheme "-ING" in speakers from California *versus* Southern America (see also Sumner & Samuel, 2009; Brunellière & Soto-Faraco, 2013; Niedzielski, 1999; Goslin et al., 2012).

Most studies so far have focused on social and cross-dialectal phonetic variations, whilst social and cross-dialectal lexical variations have not received much attention. Regarding social lexical variations, a few studies have shown that the identity of the speaker (age, gender, socio-economical status) influences lexico-semantic processing during speech perception. Walker and Hay (2011) showed such an effect in an auditory lexical decision task in which participants were listening to "old words" (typical vocabulary for older speakers) and "young words" (typical vocabulary for younger speakers). Words were pronounced by an older and a younger voice. The authors observed better accuracy and faster decisions when "voice age" and "word age" matched (Walker & Hay, 2011). Similarly, Van Berkum and colleagues (2008) investigated how lexico-semantic processing is influenced by the speaker's properties. We will extensively describe this study since it was the anchor point for our experiment (investigating a similar question using the same methodology). Van Berkum and colleagues (2008) measured the brain's electrophysiological response to sentences uttered by different speakers and explored the modulation of the N400 event-related potential (ERP) component. The N400 is a negative waveform peaking in centroparietal areas of the scalp around 400 ms after stimulus onset (and later in auditory sentence comprehension, cf. Kutas et al., 1987; van Petten et al., 1999; Hagoort & Brown, 2000). The N400 component is sensitive to the ease of lexical access, since a number of studies have shown larger amplitude deflections for low compared to high frequency words in context or in isolation (e.g., Rugg, 1990; van Petten & Kutas, 1990; Dambacher et al., 2006). Moreover, the amplitude of this component is inversely correlated with the predictability and the ease of semantic integration of the current stimulus to its previous sentence context, independently of frequency (for a review, see Kutas & Federmeier, 2011). Van Berkum and colleagues (2008) investigated the effect of contextual information such as the speaker's identity in lexico-semantic integration processes. In this study, the amplitude of the N400 elicited by words that were either consistent or inconsistent with the speaker's age (or socio-economical status, or gender) was compared. Sentences like: "I like to drink some *wine* before I go to sleep" were pronounced by children and adults and participants were asked to listen to those speech fragments for comprehension. The brain's response was time-locked to the onset of the critical word "*wine*" because according to social stereotypes, drinking alcohol is less plausible in childhood than in adulthood. The authors found that the N400 component was larger for words that were inconsistent with speaker's age, suggesting that the speaker's identity (and associated stereotypes) influenced on-line lexical integration during speech comprehension. This way, integrating a word that was inconsistent with the speaker's identity was more difficult (larger N400 mean amplitude) than integrating a consistent word (even if the sentence was semantically correct in both conditions). Thus, based on this study, we know that the signature of the influence of speaker's identity on lexico-semantic processing is a modulation of the N400 ERP component.

Social phonological and lexical variation effects on speech perception can be explained in two ways (see Creel & Tumlin, 2011 for extensive discussion on this issue). First, based on the theoretical framework of exemplar-based models, it can be assumed that lexical representations are stored in long-term memory along with social representations (Hay et al., 2006; see Sumner at al., 2014 for a dual-route approach). According to this framework, lexical representations would be stored in long-term memory as detailed memory traces containing phonetic, speaker and contextual information (Johnson, 1997; Goldinger, 1998; Foulkes & Docherty, 2006, Hay et al., 2006; Creel et al., 2008). Then, people use social knowledge to categorize perceived individuals during communication, and to build predictions and inferences based on this knowledge (see Barsalou, 2012 for a review). This way, word recognition is facilitated when the word is encountered in a social context similar to the one encountered during encoding (Sumner et al., 2014; Foulkes & Docherty, 2006; Hay et al., 2006). This framework is strongly supported by Walker and Hay's (2011) results showing a speaker's social attribute effect on words out of context (auditory lexical decision task on isolated words). On the other hand, it could be that social information extracted from the speaker's voice creates some sort of semantic context constraint. Those attributes would not necessarily be stored along with lexical representations, but would still aid lexical access through some kind of semantic priming (Walker & Hay, 2011; Nygaard & Lunders, 2002; van Berkum et al., 2008; Brennan & Hanna, 2009; Horton & Gerrig, 2002, 2005). In any case, all studies and theories would agree that there is an interplay between lexical access and social/dialectal information during word recognition (Sumner et al., 2014).

We know that social lexical variations (different usage of words depending on the age, gender and socio-economical status of the speaker) influence lexico-semantic processing during speech comprehension. As far as we know, the influence of cross-dialectal variations on lexical processing still has to be explored, which is the scope of the present study. We will ask whether on-line lexical integration is modulated by the speaker's dialect (conveyed by his accent), as it is by his age, gender or socio-economical status. The question whether social lexical variation effects are generalisable to dialectal lexical variation is worth asking since social and dialectal information differ on certain aspects: social lexical variations rely on general world knowledge (social stereotypes about what a person would likely do/eat/say) while dialectal lexical variations rely purely on linguistic knowledge/experience (vocabulary associated with a dialect). In van Berkum at al.'s (2008) study, speech perception is influenced by the speaker's attributes at the *conceptual level*: voice context makes the concept of drinking alcohol or having a tattoo more or less congruent. When manipulating cross-dialectal lexical variations, speech perception will potentially be influenced by speaker's attributes at the *lexical level*: voice context makes one specific lexical entry (one of two cross-dialectal synonyms) more or less congruent. It might be that the speaker's attributes support word integration at the semantic level (a word is congruent with conceptual expectations based on the speaker's voice) but not necessarily at the lexical level. In other words, while the extant evidence suggests that a British participant listening to an American speaker will integrate *semantic* *information* linked to stereotypes of American eating habits more easily than atypical eating habits, this does not entail easier integration of specific lexical items that correspond to the same semantic concept, e.g, *cookie* (uttered more often by Americans) vs. *biscuit* (uttered more often by British speakers). Therefore, the present study will extend previous knowledge on speaker's attribute effects to the attribute of dialect (accent). Plus, it will show whether the integration facilitation effect previously reported on semantic/conceptual processing also happens at the lexical processing stage.

The second important contribution of the present study is on investigating not only the influence of speaker's attributes (hereafter called context-dependent information), but also the contribution of context-independent factors such as word frequency, and the interplay between those two variables (see predictions below for more detail).

It is worth mentioning that the present study differs from van Berkum et al.'s (2008) study in another relevant aspect. Van Berkum and colleagues deliberately used self-referential pronouns ('I', 'mine', 'we', 'our') in their critical sentences in order to maximize the speaker's attribute effects. In the discussion, they make the assumption that similar effects should be observed without using such pronouns. In the present study, speakers were uttering sentences without self-referential pronouns (see sentence examples in Table 2). Thus, finding similar effects of speaker's attributes in the present study would generalize van Berkum et al.'s results to sentences without critical self-referential pronouns.

In order to investigate cross-dialectal lexical (vocabulary) variations in on-line speech comprehension, we tested the interplay between the dialect of the speaker (inferred by his accent) and the consistency of a word with the speaker's dialect. British native listeners were exposed to sentences pronounced by British and American speakers. Each sentence contained a critical word that was part of the British (e.g., *lift*, *holiday*) or American (e.g., *elevator*, *vacation*) vocabulary[1]. ERP

[1] Word classification as part of the British or American vocabulary is based on frequency of occurrence in British and American English corpora. Based on the British National Corpus (Davies, 2004), the word *holiday*

wave mean amplitudes elicited by the critical word of each sentence were measured and analyzed in a 2x2 design: Contextual Dialect (British *versus* American speaker's accent) and Vocabulary (critical word from the British *versus* American vocabulary). The most straightforward hypothesis is that American words should elicit a larger N400 component than British words, when British listeners are exposed to British speakers (typical N400 effect elicited by lexical frequency; Rugg, 1990; van Petten & Kutas, 1990; Dambacher et al., 2006). In other words, British listeners should integrate the word "*holiday*" more easily than the word "*vacation*", when listening to their native dialect, because the frequency of occurrence of the former is higher than the frequency of the latter.

Regarding lexical access during American dialect listening, different hypotheses are proposed. (1) We can hypothesize that speaker's attributes (context-dependent information) will influence lexical integration, as has been previously shown by van Berkum and colleagues (2008) for semantic integration. In other words, word integration would be easier when the lexical item is pronounced in the consistent dialect. Thus, the N400 component would be larger for "*vacation*" than for "*holiday*" when listening to British speakers, and vice versa when listening to American speakers. This result would further support exemplar-based models of speech perception and lexico-semantic integration (Goldinger, 1998; Foulkes & Docherty, 2006; Hay et al., 2006; Barsalou, 2012). (2) On the other hand, it is not necessary that the adaptation at the conceptual and lexical levels occurs with similar mechanisms. Lexical access might be influenced both by context-dependent information

---

is more frequently encountered (receptive vocabulary) than the word *vacation*. On the other hand, based on the Corpus of Contemporary American English (Davies, 2008), the word *vacation* is more frequently encountered than the word *holiday*. Thus, the word *holiday* is considered as part of the British vocabulary and the word *vacation* as part of the American vocabulary. Note that all the words used in the study had non-null frequencies in the British National Corpus. This way, we made sure that all the critical words would be understood by the British participants (Any vocabulary effect would be due to relative differences of frequency of occurrence and not to lack of meaning knowledge). We ran an off-line test with 22 British native speakers to verify the likelihood of the critical words. Participants were visually presented with each sentence context until the critical word, and were asked to rate the likelihood of having the last word occurring in this text (independent of what might follow). They had to rate each word from 1 (very unlikely) to 5 (very likely) or by answering "D" when they did not know the critical word. Each participant was presented with each sentence ending with the British or the American word (two lists counterbalanced across participants). British words' likelihood was rated 4.21 ±.61 on average (0.00% of "D" responses). American words' likelihood was rated 3.32 ± .83 on average (0.01% of "D" responses). Thus, despite their low frequency in British vocabulary, American words were known and judged fairly likely to occur in our sentence contexts.

(speaker’s attributes) and context-independent lexical frequency when listening to dialectal accented speech (unfamiliar accent). In this case, British listeners would easily integrate “*holiday*” in American contexts because this word is the word they encounter in their everyday experience with language (context-independent lexical frequency), but “*vacation*” would also be easily integrated because of its specific occurrence in the American context (context-dependent effect). Thus, no N400 effect should be observed when comparing British and American words during American dialect listening; the N400 effect should be reserved for the more familiar British dialect, where *“vacation”* is supported neither by context-dependent or independent information. This result would be in line with previous research on cross-dialectal phonetic/phonological variation (see Brunellière & Soto-Faraco, 2013, Loudermilk, 2013a, 2013b). Brunellière and Soto-Faraco (2013) tested whether listener’s phonological expectations are formed in accordance with the dialectal accent of the speaker (Eastern *versus* Western Catalan accents). An effect of unexpected, inconsistent, phonology was observed when the speaker used the accent familiar to the listener, but no evidence of phonological expectation was found when the speaker’s accent was unfamiliar to the listener. A similarly asymmetrical effect was reported by Loudermilk and colleagues (2013 a, b), suggesting that these facilitation effects for phonological forms that are consistent with the speaker’s accent rely on the listener having sufficient experience with an accent as to be able to form expectations. (3) Finally, we cannot discard a third alternative hypothesis, being that listeners are not influenced by the speaker's attributes (context-dependent information), and that lexical integration is mainly driven by context-independent lexical frequency. If the easiest word to integrate is the most frequent one given a specific sentence context (irrespective of the speaker's accent), then we should observe an increased N400 component for the less frequent compared to the most frequent word, in the vocabulary of the listener. In that case, the N400 component would be larger for “*vacation*” than for “*holiday*”, in both British and American dialect contexts.

Materials and Methods

Participants

Data were collected from 45 native British participants (mean age = 21 ±5 years old; 32 females; 33 right handed) from the University of Dundee. While they were living in a British speaking country, their self-rated exposure to spoken American on a 7-point scale (1: never; 7: everyday) was 4.8 (*SD* =1.6). Five subjects were excluded from the analyses due to high levels of artifactual contamination (exclusion based on blind visual inspection of EEG recordings preceding ERP processing). The volunteers participated for class credit or payment after providing informed consent. All methods were approved by the ethics committee of the University of Dundee.

Materials

To create our materials we adapted a pseudo-random selection of English sentences and vignettes from two corpora (the British National Corpus; Davies, 2004 and the Corpus of Contemporary American English; Davies, 2008). The corpora were searched for instances of our critical words. These sentences were then inspected and a target sentence was selected if it met the following criteria: 1) the meaning of the critical word was sufficiently constrained and unambiguous, 2) the predictability of the word was low given the context, 3) the critical word was neither the first nor the last word of the sentence. In some cases, surrounding sentences were included, to provide additional context. In other cases, bigrams were searched for, .e.g in cases where the ambiguity relied on a collocation such as 'fish sticks'/'fish fingers'. The experimental stimuli consisted of 76 sentence pairs across two conditions: British critical word *versus* American critical word. The sentences within each pair were identical except for the critical word, which was never the last word of the sentence (word position: M = 23rd [*SD* 5]; number of words per sentence: M = 33 [*SD* 7]). British words (e.g. *lift*, *holiday*) were defined as those that are more frequent than corresponding American synonyms in the British National Corpus (Davies, 2004). Conversely, American words (e.g. *elevator*, *vacation*) are more frequent than their British counterparts in the Corpus of

Contemporary American English (Davies, 2008). Critical words' properties are described in Table 1[2], and examples of sentences are reported in Table 2.

Sentence contexts preceding critical words were designed to avoid semantic ambiguity at the critical word. Additionally, we selected our sentential contexts in such a way that the critical word was not highly predictable. The cloze probability, defined as the predictability of the critical word given the context frame, was assessed using a completion test administered to 20 participants selected with the same inclusion criteria used for the ERP task. The task directed participants to write down the word that “first comes to mind” given the written context. The sentences were continued in 12% (SEM=2.6) of the cases with the British word and in 2% (SEM=0.6) of the cases with the American word; $t(74)=4.54$, $p<0.001$. The difference between the cloze probability of the two words was positively correlated with the difference in word frequency from the British National Corpus (r [74] =0.32, $p<0.01$) so that the more dominant a word was in British, the larger its cloze probability.

Similar cloze probabilities were obtained in an auditory sentence completion test. Each sentence, spoken by the American and the British speaker, was edited to terminate just before the critical word. These sentence fragments were presented to 23 British participants, who were directed to complete the sentences with the first word or phrase that came to mind. Each participant heard one version of each sentence fragment. Sentence versions, British and American, were counterbalanced across two presentation lists. Sentences were completed with the British target word in 12% (SEM=5.7) of trials and with the American critical word in 4% (SEM=3.0) of trials. Additionally, participants were tended to generate dialect-consistent words, with the rate of British words increasing when the sentence was spoken by a British speaker and American words

---

[2] Critical word durations were submitted to a 2 (British-accent, American-accent) x 2 (British-word, American-word) ANOVA. There was a main effect of Accent ($F[1, 75] = 20.34$, $p <.001$), showing that the American speaker was uttering critical words slower than the British speaker. This main effect highlighted the main difference between speakers, the core manipulation of the study. Importantly, there was no effect of Vocabulary ($F[1, 75] = 2.14$, $p = .15$) and no interaction ($F[1, 75] = .19$, $p = .67$). Thus, any Accent x Vocabulary interaction in the ERP pattern could not be explained by differences in the durations of the experimental material.

increasing when the sentence was spoken by an American speaker ($F[1, 22] = 14.15$, $p=.001$). This interaction reflects a marginally significant difference in the proportion of British words in response to British compared to American sentence fragments (13 ±7% vs. 10 ±4%; $t(22) = 1.85$, $p=.079$) and a significant preference for American words in response to American compared to British sentence fragments (3 ±2% vs. 5 ±4%; $t(22) = 2.6$, $p=.018$).

Thus, cloze probability was kept low in order to avoid specific lexical anticipation. Still, sentence contexts were selected in such a way that the semantic field of the context was restricted. This way, erroneous interpretation of polysemous words was avoided. For instance, in the sentence context: “She ran back into the corridor. Behind her she heard Bodie's voice. He was racing up the stairs. She entered the lift…” it is unlikely that the wrong meaning of the word “lift” would be selected.

Each sentence of a pair was recorded by two different male speakers. For the 'British-accent' condition the speaker was a British native[3], whilst for the 'American-accent' condition the speaker was a North-American native speaker of English. 'British-accent' and 'American-accent' recordings were matched on speech rate measured in terms of number of words per second (Native: $M = 2.7$ [*SD* 0.2]; Non-native: $M = 2.6$ [*SD* 0.2]; $t(151)<1$, $p=0.3$). In the 'British-accent British-word' condition, the British speaker recorded the sentence with the British critical word. In the 'British-accent American-word' condition, the speaker recorded the same sentence but uttering the American critical word. The American speaker recorded the same stimuli; sentences containing the British critical word constituted the 'American-accent British-word' condition, whereas sentences containing the American critical word constituted the 'American-accent American-word' condition. During stimulus recording, speakers were asked to produce each sentence twice (in a randomized order). Three naive raters were presented with those sentences. They were asked to choose the version that sounded "the most natural". This version was then used in the experiment.

[3] The British native speaker was from the south of England. His accent was close to the 'Received Pronunciation', and different to the one of the participant sample (predominantly Scottish and Northern English accents).

Four lists of 76 sentences (19 for each of the 4 conditions) were created and each participant was randomly assigned to one of them. In each list, the 19 sentences of each of the 4 conditions were randomly presented. The lists were created in such a way that no participant heard the same sentence in more than one condition, each variant was heard by an equal number of participants, and the longest consecutive sequence of trials of the same type was two. Among the 40 participants included in the analyses, 9 had been assigned list 1, 11 list 2, 11 list 3 and 9 list 4.

Procedure

Electrode application and total time-on-task took one hour. Subjects were instructed to listen to and comprehend 76 sentences presented over loudspeakers. The presentation of audio stimuli was programmed using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). The trials were presented in 9 blocks, separated by rest periods. Each trial began with a fixation cross centered on the screen. After 1 sec, the auditory sentence was played. The cross remained on the screen until 500 ms after sentence offset, and was followed by a 3 seconds inter-trial blank screen interval.

EEG recording and processing

A 34 channel active electrode system was connected to a recording computer utilizing BioSemi – Activision software (Biosemi Inc., Amsterdam,The Netherlands). Recordings were carried out using a BioSemi CHA-01 at a sampling rate of 2048 Hz. We used 30 electrodes mounted in an elastic cap placed according to the 10–20 system at scalp sites: AF3, AF4, F7, F8, F3, F4, Fz, FC1, FC2, FC5, FC6, T7, T8, C3, C4, Cz, CP1, CP2, CP5, CP6, P3, P4, Pz, P7, P8, PO3, PO4, O1, O2 and Oz. Additionally, electrodes were positioned on the left and right mastoids and below and at the outer canthi of the left eye to monitor vertical and horizontal eye movements. Recordings were re-referenced offline to linked mastoids using PolyRex 1.2 (Kayser & Tenke, 2003).

Data were analyzed using Brain Vision Analyser software (BrainProducts, Munich, Germany). The EEG signal was filtered off-line using a 0.03-30 Hz band-pass filter. Eye blink artifacts were corrected using the Gratton et al.'s procedure (1988). EEG epochs ranged from -200 ms to 1200 ms, time-locking the onset of the critical word of each sentence at 0 ms. We applied automated artifact rejection procedures constituted by three criteria: an amplitude criterion of ±150μV, a gradient criterion (the maximum allowed voltage step between consecutive data points) of 70μV, and a difference criterion (the maximum difference allowed between two values within 200 ms) of 100μV[4]. Baseline correction was performed with reference to pre-stimulus activity (-200 to 0 ms). Segments were averaged for each participant, each experimental condition and each electrode. The N400 mean amplitudes were measured in the standard [300-600] ms time-window following the onset of the critical word. Since it has been shown that the N400 component lasts longer in the auditory than the visual modality (see Kutas et al., 1987; van Petten et al., 1999; Hagoort & Brown, 2000), we also analyzed ERP amplitudes in the later [700-900] ms time-window. ERPs were analyzed over the entire scalp divided into 4 regions of interest (ROI; the mean amplitude measure was collapsed across the electrodes in each ROI). The 'Left Anterior' ROI included AF3, F3, F7, C3, FC5, the 'Right Anterior' ROI included AF4, F4, F8, C4, FC6, the 'Left Posterior' ROI included CP1, CP5, P3, P7, PO3 and the 'Right Posterior' ROI included CP2, CP6, P4, P8, PO4.

Statistical analyses

A repeated measures analysis of variance (ANOVA) was performed on the ERP mean amplitudes (early and late time-window) with factors: Contextual Dialect (2: British-accent, American-accent), Vocabulary (2: British-word, American-word), Anteriority (2: Anterior, Posterior) and Laterality (Left, Right). As the Mauchly tests did not indicate sphericity violations, no Greenhouse-Geisser

[4] Based on this procedure, artifactual data points (and not entire epochs) were rejected from data analysis. After artifact rejection, 6.3 % (SD = 6.8) of the data points were rejected in the British-accent British-word condition. 6.8% (SD = 7.3) were rejected in the British-accent American-word condition, 6.2% (SD = 6.9) in the American-accent American-word condition and 6.4% (SD = 6.7) in the American-accent British-word condition. The percentage of rejected data points did not significantly differ between conditions (Main effect of Contextual Dialect: $p>.3$; Main effect of Vocabulary: $p>.2$; Interaction: $p>.5$).

corrections were used. Post-hoc tests are two-tailed paired-sample tests using Bonferroni-adjusted alpha values to correct for multiple comparisons.

Results

The Grand averaged ERP results and plot of mean amplitude measures can be seen in Figures 1, 2 and 3.

[300-600] ms time-window

The repeated measures ANOVA revealed no effect of Contextual Dialect ($F[1,39] < .001$, $p = .99$), no effect of Vocabulary ($F[1,39] = .005$, $p = .94$), no effect of Anteriority ($F[1,39] = .33$, $p = .57$) and no effect of Laterality ($F[1,39] = .74$, $p = .39$). The Anteriority x Laterality interaction was significant ($F[1,39] = 5.78$, $p = .02$) but the post-hoc analysis did not reveal any significant effect (all ps > .16). The other interactions did not reach significance (all ps >.05).

The important outcome for the study was that there was no effect of Contextual Dialect, Vocabulary and no interaction between the two factors in the [300-600] ms time-window (see Figure 2).

[700-900] ms time-window

The repeated measures ANOVA did not reveal significant effects of Contextual Dialect ($F[1,39] = .11$, $p = .75$) or Vocabulary ($F[1,39] = .02$, $p=.89$). There was a significant Contextual Dialect x Vocabulary interaction ($F[1,39] = 14.94$, $p < .001$). The post-hoc analysis of the Contextual Dialect x Vocabulary interaction revealed that British words elicited a larger negative component than American words in the American-accent condition (when listening to American speakers; $p = .04$). There was a tendency for American words eliciting larger negative component than British words in the British-accent condition (when listening to British speakers; $p = .07$). There was a significant effect of Anteriority ($F[1,39] = 7.69$, $p=.008$) but no effect of Laterality ($F[1,39] = .51$, $p=.48$). The

Anteriority x Laterality interaction was significant ($F[1,39] = 8.68$, $p=.005$) revealing an effect of Anteriority in the left hemisphere ($p < .001$) but not in the right hemisphere ($p = 1.00$): ERP mean amplitudes were larger over posterior than anterior region in the left hemisphere. No other interaction reached significance (all ps > .05). Licensed by the main effect of Anteriority, we performed two separate ANOVAs in the anterior and posterior regions. Based on previous literature, the ERP modulation was expected to be larger over posterior compared to anterior sites (van Berkum et al., 2008; Kutas & Federmeier, 2011). The ANOVA factors were: Contextual Dialect (2: British-accent, American-accent), Vocabulary (2: British-word, American-word) and Laterality (Left, Right).

The ANOVA performed in the anterior region revealed a significant Laterality effect ($F[1,39] = 5.77$, $p=.02$) showing that the ERP waves were more negative over the left hemisphere. The Contextual Dialect x Vocabulary interaction was significant ($F[1,39] = 9.23$, $p=.004$) but the post-hoc analysis did not reveal any significant effect (all ps > .12). No other main effect or interaction reached significance (all ps > .05).

The ANOVA performed in the posterior region revealed a significant Laterality effect ($F[1,39] = 4.89$, $p=.03$) showing that the ERP waves were more negative over the right hemisphere. The Contextual Dialect x Vocabulary interaction was significant ($F[1,39] = 16.88$, $p < .001$). No other main effect or interaction reached significance (all ps > .05). The post-hoc analysis of the Contextual Dialect x Vocabulary interaction revealed that American words elicited a larger negative component than British words in the British-accent condition (when listening to British speakers; $p = .03$). Conversely, British words elicited a larger negative component than American words in the American-accent condition (when listening to American speakers; $p = .04$). Moreover, the electrophysiological response did not differ in the consistent or inconsistent contexts: The ERP mean amplitude did not differ between 'British accent - British word' and 'American accent -

American word' (p = 1.00). It also did not differ between 'British accent - American word' and 'American accent - British word' (p = .08).

Thus, in the [700-900] ms time-window, the processing of British words was easier when listening to British speakers and processing of American words was easier when listening to American speakers (see Figure 3).

The absence of an effect in the [300-600] ms time-window could be due to the fact that critical words were ~ 590 ms-long. In fact, critical words were 6-7 phoneme-long, which implies that they had a late uniqueness point (averaged uniqueness point for 6-7 phoneme-long words is around 5; see Luce, 1986, Figure 3). Critically, cross-dialect pairs were not matched on length, so some British words were longer than their American counterparts and vise versa. This variability in the stimuli combined with late uniqueness points (on average) and general low cloze probability may have contributed to the late emergence of our N400 effect. To test this hypothesis, we performed a further analysis designed to reduce the influence of word length. We pooled the data across the 'British accent – British word' and 'American accent – American word' conditions to create a new 'Consistent condition' and we did the same for the 'British accent – American word' and 'American accent – British word', creating a new 'Inconsistent condition'. These two new conditions contained the same words, which means they were matched for word length and power was increased[5]). When comparing Consistent and Inconsistent words, we had the clear hypothesis that the latter should elicit a larger N400 component than the former. Thus, we performed a 1-tailed t-test comparing Consistent and Inconsistent words in the [300-600] ms time-window. The N400 mean amplitude was significantly larger for inconsistent words (mean = -1.84 ±2.2 μV) than consistent words (mean = -1.11 ±2.1 μV; p = .017) over the posterior region of the scalp.

[5] Note however that this analysis reduced statistical constraints, facilitating the observation of significant effects. The results should then be considered as informative but no strong conclusion will be drawn based on this last analysis.

One potential caveat was detected in material recording: It might be that British critical words are less intelligible when uttered by American speakers than by British speakers (and similarly for American critical words uttered by British *versus* American speakers). Thus, variations in critical word integration in British *versus* American speech might be due to intelligibility and not to co-integration of a speaker's characteristics. This potential caveat was resolved by running an off-line post-test with 9 native British participants who did not take part of the ERP experiment, and who had similar social and educational history as the ERP experiment participants (students from the University of Dundee). Participants were presented with isolated words in the auditory modality. Isolated words were all the critical words of the main experiment, spoken in isolation to ensure that the intelligibility of each word per se would be rated, with no influence of the surrounding context. There were 4 blocks of trials with breaks in between. Only one version of a trial occurred in each block (British word – British accent, British word – American accent, American word – British accent, American word – American accent). Trials within each block were fully randomized for each participant. Participants were asked to rate the intelligibility of each auditory word, on a 5-point scale (1 = unintelligible; 5 = perfectly intelligible). If critical words uttered inconsistently (British words uttered by American speakers and conversely) were as intelligible as critical words uttered consistently (British/American words uttered by British/American speakers, respectively), their rating in the post-test should not differ. In fact, averaged ratings for consistent and inconsistent critical words did not differ (Consistent words, $R = 4.5 \pm .3$; Inconsistent words, $R = 4.5 \pm .3$; t test: $p = .66$). Intelligibility ratings for British critical words did not differ when uttered by the British or the American speaker ($p = .12$). Intelligibility ratings for American critical words did not differ when uttered by the British or the American speaker ($p = 1.00$). Since ratings of critical words did not differ when uttered by a consistent *versus* inconsistent speaker, we can reasonably argue that the intelligibility of critical words did not affect sentence processing (even if word recognition might vary in isolation versus sentence context).

Discussion

The current experiment investigated the electrophysiological response of the brain to cross-dialectal lexical variation in native and non-native dialects to shed light on how integrative processes occur during on-line speech comprehension. The experimental logic was that if listeners are flexible enough to take into account speaker's attributes (dialectal accent) during the incremental comprehension of sentences, words that are inconsistent with the speaker's dialect would elicit larger N400 effects than consistent words (larger N400 mean amplitude for American than British words uttered by British speakers and vice versa). On the other hand, if listeners do not integrate their knowledge about the speaker's dialect, those words that are less frequent in their own experience, e.g., those not typically used in their native dialect, would elicit larger N400 effects regardless of the dialect of the speaker (larger N400 mean amplitude for American than British words, whatever the speaker's accent). Finally, based on previous research on sociophonetics, it might be that lexical integration is mixed when listening to an unfamiliar dialect. In that case, less frequent words should elicit larger N400 effects in the familiar (native) dialect only (larger N400 mean amplitude for American than British words in British speech listening, no N400 modulation when listening to American speech).

We found that words inconsistent with the dialect of the speaker (British words uttered by American speakers and American words uttered by British speakers) elicited larger negative deflections than consistent words, mainly over posterior regions of the scalp, between 700 and 900ms after word onset. This result suggests that listeners integrate indexical sources of information - such as the speaker's accent – during lexical access. These results support the first alternative hypothesis, revealing a large flexibility to adapt to the speaker's attributes. Contrary to the second alternative hypothesis, it seems that context-independent lexical frequency of the word-form (word frequency in the vocabulary of the listener) is not the main factor modulating lexical processing during sentence comprehension. Context-dependent information (such as the speaker's dialectal accent) is mainly driving lexical processing (cf. below for further discussion of the consequences of this observation). Finally, the results do not mimic the ones previously reported in similar studies on phonological variations. Thus, we can conclude that adaptation to the speaker's phonological and lexical characteristics does not follow the same rules. While listeners are sensitive to a speaker's phonological consistency with their own accent or dialect (Brunellière & Soto-Faraco, 2013, Loudermilk, 2013a, 2013b), they seem to be sensitive to a speaker's vocabulary consistency both in their own and less familiar dialects. Note, however, that sentence contexts used in Brunellière and Soto-Faraco (2013)'s study were highly semantically constrained which was not the case in the present study.

Additionally, we did not find differences between consistent words in the native *versus* the non-native dialect: Integration of American words in American accented speech was similar to the integration of British words in British accented speech. This result is consistent with previous studies showing that lexical integration is not disrupted by dialectal accents when words are consistent with the vocabulary linked to these accents. Goslin and colleagues (2012) showed, for instance, that the N400 mean amplitude did not differ between native and regional dialectal accents when participants were presented with words that were consistent with both accents (e.g., "Roger searched the church tower for the pastor" where "pastor" is the critical word, belonging to the common vocabulary of both types of accents; see also Sumner, 2011)[6]. The present study confirms those previous results, and goes beyond by showing that lexical integration is not influenced by familiar regional accents, as long as the words to be integrated are consistent with the accent's vocabulary. In fact, a word that is not the most common in the speaker's vocabulary is more difficult to integrate in the sentence context.

[6] Note that lexical integration (eliciting the N400 component) being independent of dialectal accent does not mean that dialectal accent does not affect word processing at all. Goslin and colleagues (2012) observed, in fact, PMN (Phonological Mapping Negativity) amplitude modulations between native and dialectal accents, reflecting processing differences at the earliest stages of spoken word recognition.

The present results are in line with previous research showing that semantic processing is influenced by speaker's attributes (context-dependent information) such as gender, age or socio-economical status (van Berkum et al., 2008; Walker & Hay, 2011). The first contribution of the present study is to extend these previous results to lexical access and accented speech: Participants use speaker's characteristics not only to predict a specific concept (what the speaker is likely to say/eat/like) but also a specific lexical entry (which word the speaker is likely to use). Furthermore, this effect is not only observed for social characteristics such as gender or age, but also cross-dialectal characteristics such as dialectal accent. Our second contribution is that we showed not only that speaker's attributes influence lexical processing, but also that this influence is larger than context-independent lexical frequency. In other words, a lexical item embedded in a sentence is integrated better when it fits with the speaker's accent, independently of its frequency of occurrence in the receptive vocabulary of the listener. When listening to American speakers, British listeners integrate American better than British words, even if the frequencies of the former are much lower than the latter (see below for methodological implications). Thirdly, as expected by van Berkum and colleagues (2008), usage of self-referential pronouns is not mandatory to observe speaker's attribute effects. Such effects were observed in the present study, using sentences without critical self-referential pronouns. Finally, our results show that speaker's attributes influence lexical integration even with no constrained sentence context. The effect we observed cannot be explained by word anticipation based on the speaker's attributes, since sentences were very low-constrained (low cloze probability).

As it was the case for social lexico-semantic variation effects on speech perception, the reported effects of cross-dialectal lexical variation can be explained in two ways. The exemplar-based models of language perception stipulate that utterances are stored in long-term memory as separate exemplars, complete with acoustic/phonetic detail (see Foulkes & Docherty, 2006; Goldinger, 1998) and social information about the speaker (Hay et al., 2006; Creel et al., 2008; Sumner et al., 2014; Barsalou, 2012). Our results are consistent with this framework, by showing that speaker's attributes such as dialectal accent affect speech processing. It might be that British and American words are stored in long-term memory along with speakers' dialectal accents (see Walker & Hay, 2011 for such an assumption for speaker's age). The word *holiday*, encountered predominantly from British speakers, would have a distribution of predominantly British-accented exemplars and the word *vacation* would have a distribution of predominantly American-accented exemplars. Any utterance of the word *holiday* would then be easier to process when pronounced with a British accent, and *vacation* would be easier to process when uttered by an American speaker, due to the accent-match between encoding and retrieval (Walker & Hay, 2011). Nevertheless, it could be that lexical entries are not stored along with speaker's accent, but that the facilitation effect is due to some sort of semantic priming. Speaker's attributes could be extracted from speech processing and aid lexical access through semantic priming, without being necessarily stored along with lexical representations (Walker & Hay, 2011; Nygaard & Lunders, 2002; van Berkum et al., 2008; Brennan & Hanna, 2009; Horton & Gerrig, 2002, 2005). Based on this study, we cannot support one of those alternative assumptions. Still, previous research showed that acoustic-match encoding (facilitation for same speaker's attributes during encoding and retrieval) is automatic and talker-semantic encoding (facilitation for semantic priming induced by speaker's attributes) is not. Talker-semantic encoding seems to be sensitive to task utility, in the sense that it influences speech perception only when it is highly relevant (see Creel & Tumlin, 2011). We can speculate that talker's accent encoding is highly relevant in the current task, where participants listen to familiar and unfamiliar accented speech. Thus, it is very likely that talker-semantic encoding played a crucial role in the present pattern of results, as in previous studies (Creel & Tumlin, 2011; van Berkum et al., 2008; Staum Casasanto, 2008). Still, we cannot disregard the very likely possibility that an interplay between both talker-semantic and acoustic-match encoding was the core mechanism at play (Creel & Tumlin, 2011).

Another important theoretical implication of our findings is that phonological and lexico-semantic integration during accented speech comprehension do not seem to follow the same principles. We argue that the interplay between context-dependent and context-independent information varies according to the level of speech processing (at least in the context of dialect accented speech). During dialectal speech perception, phonological properties extraction might be based on a mix between general likelihood (context-independent lexical frequency of the word-form) and local probability (frequency of the word-form in the regional accent context; Brunellière & Soto-Faraco, 2013). The large influence of context-independent information would be explained by the fact that phonological rules are regular, which makes a word's phonological form highly predictable based on general likelihood. On the other hand, later processing of lexical and semantic integration might depend more on context-specific information (speaker's attributes), because of the idiosyncratic nature of lexical variations in language use. Thus, lexical integration in the same context of dialectal speech perception would be mainly based on context-dependent information.

A potential caveat when interpreting our findings is that the effects expected in the N400 time-window were observed in the [700-900] ms time-window. However, we have reasons to believe that these effects reveal a late N400 effect. First, previous studies on sentence processing have reported late and long-lasting N400 effects in the auditory modality (see for instance Kutas et al., 1987; van Petten et al., 1999; Hagoort & Brown, 2000). Moreover, the fact that the effect was late is not surprising given that (1) the sentences were not constrained at all, which means that no anticipation of the critical word was possible (which would have sped up word integration and elicited earlier N400 effects). (2) The alternative inconsistent word (British word in American context and vice versa) did not elicit a semantic violation; it did not even modify the main sentence understanding in the sense that it was a sort of synonym (inter-dialect synonym). The N400 effect is usually observed in an earlier time-window when the alternative critical word affects the general meaning of the sentence (typical N400 effect elicited by semantic violations; Kutas et al., 1987; Kutas & Federmeier, 2011). (3) Finally, the critical words in our experiment were 590 ms-long on average. Thus, it is not surprising that the N400 effect appears late in time, since the uniqueness point of each critical word has to be reached before any lexical effect (N400 effect) can be observed. This argument was supported by the observation of a consistency effect in the [300-600] ms time-window. When comparing Consistent and Inconsistent words, we reduced the noise in the ERP data due to variability in the uniqueness points. The consistency effect was significant in the typical [300-600] ms time-window of the N400 component, which strongly suggests that the Contextual Dialect x Vocabulary interaction reaches significance in a late time-window because of the long length of the critical words.

For all those reasons, we argue that the effect reported in the [700-900] ms time-window is a late N400 effect[7]. To date, the mechanisms and functional significance of the N400 component are still subjects of debate. While some authors argue that the key mechanism underlying the N400 component is the ease of integration of a word in its context (Hagoort & van Berkum, 2007), other researchers have proposed that the component is modulated by the contextually-induced preactivation of words and its facilitatory effect on word recognition (for reviews, see Lau et al., 2008; Federmeier et al., 2007). While it is unlikely that a single cognitive process underpins the N400, we interpret our results under the integration view because the sentential contexts used in this experiment did not afford high degrees of lexical anticipation. We specifically designed our stimuli not to highly constrain the lexical options at the point of the critical words, a design characteristic which was validated by the low cloze probability (12%) of critical words. Thus, whether we accept or reject the view that the electrophysiological response recorded in the [700-900] ms time-window reflects word integration, we can conclude that the integration of a word in its context is modulated by indexical information such as the speaker's accent. Crucially, our findings do not depend on an interpretation in terms of electrophysiological components.

[7] Another argument in favor of a late N400 effect was that alternative interpretation in terms of later positive components was not possible. Our data cannot be interpreted in terms of P600 (Osterhout, 1997; Friederici, 2002), Late Positivity (LP; Thornhill and Van Petten, 2012; Van Petten and Luka, 2012) or Late Positive Component (LPC; Moreno et al., 2002) modulation, since inconsistent critical words should elicit larger positivities, and the opposite was observed in our study.

The observation that word integration does not only depend on its frequency of occurrence, but also (and even more importantly) on the context-dependent (indexical) information extracted from the context has also important methodological implications. The word '*holiday*' (most frequent for British listeners) is easier to integrate than the word '*vacation*' when uttered by a British speaker, but it is more difficult to integrate when the speaker has an American accent. Thus, context-dependent information (frequency of the word-form in the dialect of the speaker) better explains lexical integration in speech comprehension than context-independent lexical frequency. As a consequence, studies investigating auditory language comprehension should control for lexical frequency in the dialect of the speaker, and not the dialect of the listener (participant tested in the experiment). In other words, the variable of lexical frequency which is widely used in experimental psychology and linguistics should be considered a valid variable only when listeners are processing language in the same accent as their own, and not in a different regional/dialectal accent. This observation highlights the important role of the frequency statistics from conversational speech in the development of spoken word recognition theories (see Connine et al., 2008 for similar conclusions on variant form frequency in spoken word recognition).

This first study on vocabulary (cross-dialectal lexical variation) integration in accented speech calls for further investigation. First, it would be interesting to investigate such effect not only in dialectal but also foreign accents. We can hypothesize that when listening to foreigners speaking in English, word integration would mainly be driven by context-independent lexical frequency, since no precise inference can be made on particular vocabulary usage. In that case, the N400 component would be larger for "*vacation*" than for "*holiday*", in both British and foreign accent contexts. Second, we can safely assume that the modulation of the N400 component goes along a continuum based on the relative frequencies of the synonyms in each language, and based on the familiarity of the listener with the non-native dialect. We hypothesize that, at a trial-by-trial level, the larger the difference in frequency between British and American English, the larger the N400 modulation. We can also assume that, at the individual level, the more familiar a listener is with the non-native dialectal accent, the larger the N400 incongruity effect should be when listening to an American speaker using a British word (see Walker & Hay, 2011 for similar expectations with different exposure to older *versus* younger speakers).

Conclusion

We have shown that knowledge about the speaker's dialect and the common vocabulary of the dialect are relevant for on-line sentence processing. British natives listening to British speakers integrate more easily the British word '*holiday*' than the American word '*vacation*' because of its higher frequency in their vocabulary. On the other hand, the word '*vacation*' is easier to integrate than the word '*holiday*' when listeners are exposed to American accented speech. These results show that the speaker's accent influences lexical processing, both in the familiar and less familiar accent conditions. Moreover, these results have important methodological implications, revealing that context-dependent word frequency better explains word recognition than context-independent frequency.

Acknowledgements

This research was funded by grants from the Spanish Government (PSI2011-23033 and Consolider Ingenio 2010 CSD2007-00048), from the Catalan Government (SGR 2009-1521) and from the European Research Council under the European Community's Seventh Framework (FP7/2007-2013 Cooperation grant agreement 613465-AThEME). CM was supported by the Basque Foundation for Science (IKERBASQUE) and the BCBL institution.

References

Barsalou, L. W. (2012). The human conceptual system, in The Cambridge Handbook of Psycholinguistics, eds M. Spivey, K. McRae and M. Joanisse (New York: Cambridge University Press), 239–258.

Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialog. Topics in Cognitive Science, 1(2), 274–291.

Brunellière, A., Soto-Faraco, S. (2013). The speakers' accent shapes the listeners' phonological predictions during speech perception. Brain and Language, 125, 82-93.

Campbell-Kibler, K. (2010). Sociolinguistics and perception. Language and Linguistics Compass, 4, 377–389.

Connine, C. M., Ranbom, L. J., & Patterson, D. J. (2008). Processing variant forms in spoken word recognition: The role of variant frequency. Perception & Psychophysics, 70(3), 403–411.

Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2008). Heading the voice of experience: The role of talker variation in lexical access. Cognition, 108, 633–664.

Creel, S. C., Bregman, M. R. (2011). How Talker Identity Relates to Language Processing. Language and Linguistics Compass, 5, 190–204.

Creel, S. C., & Tumlin, M. A. (2011). On-line acoustic and semantic interpretation of talker information. Journal of Memory and Language, 65, 264–285.

Creel, S. C. (2012). Phonological similarity and mutual exclusivity: on-line recognition of atypical pronunciations in 3-5-year-olds. *Developmental Science*, *15*(5), 697–713.

Dambacher, M., Kliegl, R., Hofmann, M., Jacobs, A.M. (2006). Frequency and predictability effects on event-related potentials during reading. Brain Research, 1084(1), 9-103.

Davies, M. (2004) *BYU-BNC*. Based on the British National Corpus from Oxford University Press.

Davies, M. (2008) *The Corpus of Contemporary American English: 450 million words, 1990-present*.

Drager, K. (2010). Sociophonetic variation in speech perception. Lang. Ling. Comp. 4, 473–480.

Evans, B.G., Iverson, P. (2004). Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences. The Journal of the Acoustical Society of America, 115(1), 352–361.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., Kutas, M. (2007). Multiple effects of sentential constraint on word processing. Brain Research, 1146, 75–84.

Foulkes, P., Docherty, G. (2006). The social life of phonetics and phonology. Journal of Phonetics, 34, 409–438.

Friederici, A.D. (2002).Towards a neural basis of auditory sentence processing. Trends in Cognitive Sciences, 6,78-84.

Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. Psychological Review 105. 251–279.

Goslin, J., Duffy, H., Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. Brain and Language, 122, 92-102.

Gratton, G.,Coles, M.G. H., Sirevaag, E.J., Eriksen, C. W.,n Donchin, E. (1988). Pre- and post-stimulus activation of response channels: A psychophysiological analysis. Journal of Experimental Psychology: Human Perception and Performance, 14, 331–344.

Guy, G. R. (2011). Sociolinguistics and formal linguistics. In Ruth Wodak, Barbara Johnstone and Paul Kerswill (eds.) The SAGE Handbook of Sociolinguistics. London: Sage Publications, 249–264.

Hagoort, P. & Brown, C.M. (2000). ERP effects of listening to speech: semantic ERP effects. Neuropsychologia 38, 1518-1530.

Hagoort, P., van Berkum, J. J. A. (2007). Beyond the sentence given. Philosophical Transactions of the Royal Society B, 362, 801–811.

Hay, J., Warren, P., Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. Journal of Phonetics, 34, 458–484.

Horton, W. S., & Gerrig, R. J. (2002). Speakers' experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. Journal of Memory and Language, 47, 589–606.

Horton, W. S., & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. Cognition, 96, 127–142.

Horton, W.S., Slaten, D.G. (2012). Anticipating who will say what: The influence of speaker-specific memory associations on reference resolution. Mem Cognit 40: 113–126.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J.W. Mullennix (eds.), Talker Variability in Speech Processing, 145–165. San Diego: Academic Press.

Johnson, K. A., Strand, E. A., & D'imperio, M. (1999). Auditory–visual integration of talker gender in vowel perception. Journal of Phonetics, 24, 359–384.

Johnson, K. (2005). Speaker normalization in speech perception. The handbook of speech perception, ed. By D. B. Pisoni and R. Remez, 363–89. Oxford: Blackwell Publishers.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. Journal of Phonetics, 34, 485–499.

Kayser, J., Tenke, C.E. (2003). Optimizing PCAmethodology for ERP component identification and measurement: theoretical rationale and empirical evaluation. Clin. Neurophysiol, 114 (12), 2307–2325.

Kutas, M., Neville, H.J., Holcomb, P.J. (1987) A preliminary comparison of the N400 response to semantic anomalies during reading, listening, and signing. Electroencephalogr. Clin. Neurophysiol. (Suppl. 39), 325–330.

Kutas, M., Federmeier, K.D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). Annu. Rev. Psychol, 62, 621–647.

Labov, W. (1972). Sociolinguistic Patterns. University of Pennsylvania Press, Philadelphia.

Lau, E., Phillips, C., Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. Nature Reviews Neuroscience, 9(12), 920–933.

Loudermilk, B. (2013a) Cognitive Mechanisms in the Perception of Sociolinguistic Variation. PhD Dissertation. University of California.

Loudermilk, B. (2013b). Psycholinguistic Approaches, In R. Bayley, R. Cameron & C. Lucas (Eds.), The Oxford Handbook of Sociolinguistics.

Luce, P.A. (1986). A computational analysis of uniqueness points in auditory word recognition. Perception and Psychophysics, 39, 155-158.

Moreno, E. M., Federmeier, K. D., & Kutas, M. (2002). Switching languages, switching palabras (words): An electrophysiological study of code switching. Brain and Language, 80(2), 188–207

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. Journal of Language and Social Psychology, 18, 62–85.

Nygaard, L. C., & Lunders, E.A. (2002). Resolution of lexical ambiguity by emotional tone of voice. Journal of Memory and Cognition 30(4). 583–593.

Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. Brain & Language, 59, 494-522.

Remez, R. E., Rubin, P. E., Nygaard, L. C., & Howell, W. A. (1987). Perceptual normalization of vowels produced by sinusoidal voices. Journal of Experimental Psychology: Human Perception and Performance, 13, 40–61.

Rugg, M.D. (1990). Event-related brain potentials dissociate repetition effects of high and low frequency words. Memory and Cognition, 18, 367–379.

Staum Casasanto, L. (2008) Does social information influence sentence processing? In: Proceedings of the 30th Annual Conference of the Cognitive Science Society, ed. B. C. Love, K. McRae, & V. M. Sloutsky, pp. 799–804.

Strand, E. A. (1999). Uncovering the role of gender stereotypes in speech perception. Journal of Language and Social Psychology, 18, 86–99.

Sumner, M., Samuel, A.G. (2009). The effect of experience on the perception and representation of dialect variants. Journal of Memory and Language, 60, 487–501.

Sumner, M. (2011). The role of variation in the perception of accented speech. Cognition, 119(1), 131–136.

Sumner, M., Kim, S. K., King, K. B. M. (2014). The socially weighted encoding of spoken words: a dual-route approach to speech perception. Frontiers in psychology 4.

Thomas, E. R. (2002). Sociophonetic applications of speech perception experiments. American Speech, 77, 115–147.

Thornhill, D. E., and Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: frontal positivity and N400 ERP components. Int. J. Psychophysiol. 83, 382–392.

Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., Hagoort, P. (2008). The neural integration of speaker and message. Journal of Cognitive Neuroscience, 20, 580–591.

van Petten, C., Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. Brain and Cognition, 18(4), 380-393.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., Parks, M. (1999) Time course of word identification and semantic integration in spoken language. J. Exp. Psychol. Learn. Mem. Cognit. 25, 394–417.

Van Petten, C., and Luka, B.J. (2012). Prediction during language comprehension: benefits, costs, and ERP components. Int. J. Psychophysiol. 83, 176–190.

Walker, A., and Hay, J. (2011). Congruence between ‘word age’ and ‘voice age’ facilitates lexical access. Lab. Phon. 2, 219–237.

Captions

Table 1: British and American critical word properties. Averages across items are presented (standard deviations in brackets). British and American critical words significantly differed in British and American frequency (*p*s <0.001), but did not differ in acoustic duration (p = .18), length in phonemes (.23) and number of phonological neighbors (.41).

Table 2: Examples of speech fragments used in the experiment. Critical words are in italic (*British/American* word).

Figure 1: Grand averaged ERPs over the 20 electrodes included in the analyses (AF3, AF4, F7, F3, F4, F8, FC5, C3, C4, FC6, CP5, CP1, CP2, CP6, P7, P3, P4, P8, PO3, PO4). **Up:** Grand averages for British listeners exposed to British speakers uttering a British (black line) or American (grey line) critical word. **Down:** Grand averages for British listeners exposed to American speakers uttering an American (black line) or British (grey line) critical word. Time zero corresponds to the onset of the critical word of the sentence. Negativity is plotted up.

Figure 2: Bar-plots of the mean amplitudes for each condition in each region of interest (ROI) in the [300-600] ms time-window. **a.** ERP mean amplitude in the left anterior ROI (AF3, F7, F3, FC5, C3). **b.** ERP mean amplitude in the right anterior ROI (AF4, F4, F8, C4, FC6). **c.** ERP mean amplitude in the left posterior ROI (CP5, CP1, P7, P3, PO3). **d.** ERP mean amplitude in the right posterior ROI (CP2, CP6, P4, P8, PO4). Error bars depict standard errors. AaAw = American accent - American word; AaBw = American accent - British word; BaAw = British accent - American word; BaBw = British accent - British word.

Figure 3: Bar-plots of the mean amplitudes for each condition in each region of interest (ROI) in the [700-900] ms time-window. **a.** ERP mean amplitude in the left anterior ROI (AF3, F7, F3, FC5, C3). **b.** ERP mean amplitude in the right anterior ROI (AF4, F4, F8, C4, FC6). **c.** ERP mean amplitude in the left posterior ROI (CP5, CP1, P7, P3, PO3). **d.** ERP mean amplitude in the right posterior ROI (CP2, CP6, P4, P8, PO4). Error bars depict standard errors. AaAw = American accent - American word; AaBw = American accent - British word; BaAw = British accent - American word; BaBw = British accent - British word.